

REMARKS

Claims 1-10, 12-18 and 22-29 are currently pending in the subject application and are presently under consideration. Claims 1, 8, 15, 16, 22, and 26 have been amended as shown on pages 2-6 of the Reply.

Favorable reconsideration of the subject patent application is respectfully requested in view of the comments and amendments herein.

I. Rejection of Claim 16 Under 35 U.S.C §112

Claim 16 stands rejected under 35 U.S.C. §112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. In particular, the Office Action indicates that the phrase “or a combination thereof” is indefinite as to what comprises the combination.

Claim 16 has been amended as shown on page 4 of this Reply to remove the phrase “or a combination thereof.”

Accordingly, Applicants respectfully request withdrawal of this rejection.

II. Rejection of Claims 1-14 and 26-29 Under 35 U.S.C. §101

Claims 1-14 and 26-29 stand rejected under 35 U.S.C. §101 because the Examiner contends the claimed invention is directed to software *per se* and functional descriptive material. Claims 1, 8, and 26 have been amended to address the Examiner’s concerns in this regard. It is therefore respectfully requested that this rejection be withdrawn.

III. Rejection of Claims 1-10, 12-18 and 22-29 Under 35 U.S.C. §103(a)

Claims 1-10, 12-18 and 22-29 stand rejected under 35 U.S.C. §103(a) as being unpatentable over US Patent Application Publication Number 2002/0129014 issued to Brian S. Kim (“Kim”) and US Patent Application Number 2003/0208482 issued to Brian S. Kim (Kim-482). It is respectfully submitted that this rejection should be withdrawn for at least the following reasons. Neither Kim nor Kim-482, alone or in combination, teach or suggest each and every feature of the subject claims.

A factfinder should be aware, of course, of the distortion caused by hindsight bias and must be cautious of arguments reliant upon *ex post* reasoning. See *KSR v. Teleflex*, 550 U.S. ___, 127 S. Ct. 1727 (2007) citing *Graham v. John Deere Co. of Kansas City*, 383 U. S. 1, 36 (warning against a “temptation to read into the prior art the teachings of the invention in issue” and instructing courts to “guard against slipping into the use of hindsight”) (quoting *Monroe Auto Equipment Co. v. Heckethorn Mfg. & Supply Co.*, 332 F. 2d 406, 412 (CA6 1964))).

The subject claims relate to a web page indexing mechanism that can utilize anchor text and related reference information gleaned from pages external to the page being indexed. A web crawler can retrieve a web page to be indexed, parse the page for links to other pages, and store anchor text for the links, together with the associated destination URL information, in a page data store. Anchor text and other reference information related to the page being indexed can then be retrieved from the data store, merged with the retrieved page, and passed to an index builder for indexing. Reference information can include the anchor text of external links to the page, as well as relevant text in proximity to the link in the referencing page. Indexing pages together with this extrinsic information can create a more robust index of pages and facilitate improved search results. Additionally, and in order to maintain a clean page data store, URLs that fail to download when requested by the crawler can, after a predetermined period of time, be placed in a state that will cause the failed URL to be removed from the page data store (if the URL had been downloaded and indexed previously). In particular, amended independent claim 1 recites, *a crawler component that receives the page, retrieves the reference information associated with the page from the page data store, and provides the page and the reference information to at least an index building component; wherein **failure to receive a requested page after a first predetermined period of time causes the URL for the page to be removed from the page data store after a second predetermined period of time.***

Kim does not disclose these aspects. Kim (and its divisional, Kim-482) relates to a web page indexing system that collects pages using a web crawler and creates a data structure of indexed pages. The pages are indexed and ranked based on an analysis of keywords intrinsic to the indexed pages, as well as extrinsic keywords parsed from other pages that reference the indexed pages. However, neither Kim nor Kim-482 disclose any mechanisms for automatically

removing dead page URLs from the system, much less the removal technique disclosed in amended independent claim 1, whereby a page that fails to download after a predetermined period of time is then flagged for removal after a second predetermined period of time.

Similarly, amended independent claim 8 recites, *a parser component that parses the one or more pages for another page referenced on the one or more pages, and accumulatively stores reference information associated with the another page in a page data store, the reference information comprising descriptive information that is in proximity to anchor text associated with a referencing uniform resource locator that references the another page, **wherein failure to receive the one or more page after a first predetermined period of time causes the URL for the page to be removed from the page data store after a second period of time.*** The cited references do not disclose these aspects, as discussed *supra*.

Likewise, amended independent claim 15 recites, *deleting the information for a page from the data store when the page cannot be retrieved for a predetermined period of time*, and as already noted, Kim and Kim-482 are silent regarding this feature.

Amended independent claim 22 recites, *a parser component that parses the one or more pages for another page referenced on the one or more pages, incrementally accumulates reference information associated with the another page from each of the one or more pages when crawled, and stores such reference information in a page data store, the reference information comprising descriptive information that is in proximity to anchor text associated with a referencing uniform resource locator that references the another page, **wherein failure to receive the one or more pages after a first predetermined period of time causes the URL for the one or more pages to be removed from the page data store after a second period of time.*** As discussed above, neither cited reference discloses removal of URLs from a page data store upon failure of the page to download.

Disclosing similar features, amended independent claim 26 recites, *means for removing the page from the data store when the page cannot be received after a predetermined period of time*. Neither Kim nor Kim-482 teach or suggest this feature, as already discussed.

In view of at least the foregoing, it is respectfully submitted that Kim and Kim-482, individually or in combination, do not teach or suggest each and every feature set forth in amended independent claims 1, 8, 15, 22, and 26 (and all claims depending there from), and as

such fail to make obvious the subject invention. It is therefore requested that this rejection be withdrawn.

CONCLUSION

The present application is believed to be in condition for allowance in view of the above comments and amendments. A prompt action to such end is earnestly solicited.

In the event any fees are due in connection with this document, the Commissioner is authorized to charge those fees to Deposit Account No. 50-1063 [MSFTP512US].

Should the Examiner believe a telephone interview would be helpful to expedite favorable prosecution, the Examiner is invited to contact applicants' undersigned representative at the telephone number below.

Respectfully submitted,

AMIN, TUROCY & CALVIN, LLP

/Himanshu S. Amin/

Himanshu S. Amin

Reg. No. 40,894

AMIN, TUROCY & CALVIN, LLP
24TH Floor, National City Center
1900 E. 9TH Street
Cleveland, Ohio 44114
Telephone (216) 696-8730
Facsimile (216) 696-8731